DOCUMENT RESUME

ED 102 710 EA 006 831

AUTHOR Smock, H. Richard
TITLE The Practice of Evaluation as a Purposeful Activity.
PUB DATE Apr 75
NOTE 8p.; Paper presented at the Annual Meeting of the American Educational Research Association (60th, Washington, D.C., March 31-April 4, 1975)

EDRS PRICE MF-$0.76 HC-$1.58 PLUS POSTAGE
DESCRIPTORS Bibliographies; *Definitions; *Educational Assessment; *Educational Objectives; Educational Planning; Evaluation; Evaluation Criteria; *Evaluation Needs; Literature Reviews; *Measurement

ABSTRACT

Recent events in the world of evaluation suggest the need to reassess the basic purposes that justify its existence. In the long run, the language of evaluation must be shaped by the nature and intents of the audience and by the nature of what is being evaluated. It is the nature of what is being evaluated that must determine the methodologies, processes, and reporting formats of evaluation. To too many people, evaluation still means measurement. According to one authority, measurement basically involves the use of numerical values to represent attributes of objects; assessment includes measurement plus those judgemental activities that determine what and how to measure; and some aspects of evaluation are outside the realm of both measurement and assessment. The nature of what is being examined and the purpose of the examination determine whether one is involved in measurement, assessment, or evaluation. The dissatisfaction with defining every evaluation problem as a measurement problem is plain. The recognition that the nature of the problem has an important bearing on the choice of method is encouraging. (Author/JG)

ED102710

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

THE PRACTICE OF EVALUATION AS A PURPOSEFUL ACTIVITY*

H. Richard Smock
Office of Instructional Resources
University of Illinois Urbana-Champaign

Mountains of words have been written about what evaluation is and how evaluators might proceed with their work. Evaluation models are as prolific as makes and sizes of automobiles - and both products are touted in about the same way. Certainly the idea of a yearly changeover seems to have become firmly entrenched in the field of evaluation, and for evaluators the annual meeting of the American Educational Research Association has become the equivalent of the New York International Automotive Exhibition. And like the automobiles spewed forth by the great Detroit machines, evaluation models have grown progressively grander, progressively more complex, and progressively more expensive to operate. In an affluent society of seemingly unlimited resources, large, complex, and expensive automobiles seemed to suit our style of living. In the same way, in the days of seemingly endless government affluence encouraging the employment of evaluators, large, complex, and expensive evaluations seems to have suited the style of living to which we would all like to be accustomed. But recent events in both the world of the transportation industry and in the world of evaluation suggest that the time has come for each to reassess the basic purposes on which its existence is justified. Both have been feeding off their marketing successes for so long that the assumptions on which their existence

EA 006 831

---

*Paper presented as part of a symposium, "Metaphor in Evaluation," at the annual meeting of the American Educational Research Association, Washington, D. C., March 30 - April 3, 1975.

depends have been largely forgotten. A realization of the finite nature of our natural resources will certainly cause the transportation industry to face a larger reality than that provided by the profit motive, and an impatient public, questioning the value of many academic pursuits, brings us face to face with issues about the ends rather than the means of our evaluative efforts.

Few of us would classify ourselves as operating in the world of "basic" science. Most evaluators are quite satisfied with the notion of their work as an applied field serving an end related to the educational endeavor. Evaluation has purpose - and a practical one at that. A problem of becoming throughly engrossed in any field of study is that one is led inevitably to a preoccupation with means and, as long as others are willing to support us, we can luxuriate in the intellectual stimulation of ever more cogently defining and refining our procedures. In the process our language becomes more and more esoteric. In the process the elegance of our methodology takes precedence over the purposes of our activities. Our preoccupation with means has cost us dearly. A United States Senator recently told a group of educational researchers it was unreasonable to assume that there is support for education research and development in our country. The Senator spelled out one reason for a lack of political support.

> Educators often speak a language of their own, one which is unintelligible to the uninitiated. It is very unrealistic for the education establishment to come to the Hill, speak arcane words and phrases, and then react in an offended manner when they are not understood. When testimony is given, it must be in simple language that the generalist can conprehend, or otherwise your story will never get across (Pell, 1975).

Would anyone deny that the language of evaluation is any less arcane than that of other educational specializations? But if we are being less than successful in communicating with our benefactors are we doing better with our clients? It was Stake (1972a) who said, "My measurements are praised often by my colleagues, seldom by my clients." I don't think much has changed since then, and part of the problem relates to the language evaluators use to communicate with their audiences. Our choice of language in part relates to our lack of agreement

about the ends to be achieved by evaluation. If it is possible to agree about the ends to be served through an activity, the necessary means of attaining those ends will become more clear. I believe that in the long run the language of evaluation must be shaped by (1) the nature and intents of the audiences considering evaluative information and (2) the nature of what is being evaluated.

The Nature and Intents of the Audiences Considering Evaluative Information

Stake (1972b) has through much of his writing sensitized evaluators for the need to consider the audiences addressed by an evaluation. Scriven (1967) has discussed at length the roles and goals of evaluation, and partially relates the roles of evaluation to the purposes of the audience. Perhaps Stufflebeam (1971) has made us most keenly aware of the need for quality evaluations to have credibility, timeliness and importance. It was also Stufflebeam who offered a definition of evaluation as being the process of delineating, obtaining, and providing useful information for judging "decision alternatives." So we have had over the years authors reminding us to consider our audiences, to consider the roles as well as the goals that evaluation plays, to attend to attributes that make evaluations useful to audiences, and that the proper use of evaluation is as a tool in the decision-making process. All of these are important points. None, except Stufflebeam's suggestion that evaluations are for decision-making purposes sufficiently address the question of the uses to which audiences actually put evaluation information. Several of the purposes for which I have seen audiences use evaluations prove considerably more complex and also considerably more interesting than simply using evaluation as a tool for decision making. Allow me to list a few:

1. Evaluation as problem solver. As one source of legitimate information an evaluation often serves as a stepping stone to problem solving. Formative evaluations can be extremely influential in developmental activities. And summative evaluation has proven helpful in estimating the quality and pinpointing problems in a variety of programs. Good evaluations that are attentive to the user's need are

helpful in alleviating problems. Using evaluation reports to help alleviate problems is quite different from using them as decision makers. It's for dialogue, not decisions, that evaluations find their most constructive use.

2. Evaluation as change agent. Both the process and the product of evaluation often can act as a change agent in particular situations. Formally documenting and describing what is already part of the informal communication network can have a powerful impetus for change. If the evaluation is done by an "outsider" rather than an "insider" this documentation can have an even more profound effect.

3. Evaluation as the devil's advocate. One way that evaluations act as change agents is by allowing themsleves to be used as "the devil's advocate." In an obviously unreferenced letter from an evaluator to a client, the following sentence appeared. "Your role would be that of a recipient of an evaluator's recommendation, rather than that of chief architect of a devil's plan." Being able to point to an evaluation report as the instigator for change is a more common use of evaluation than many evaluators realize. Whoever is advocating the change can be less personally and subjectively involved.

4. Evaluation as convention. Every project needs its evaluation. And everyone knows that evaluations must be done in a certain way if they are to be respectable. House (1973) has spelled out in some detail the constraining forces of a technology of evaluation rooted in measurement and the conventional nature of evaluations emanating from that technology. He has reminded us that governmental agencies only accept "hard" data. Yet few are complimenting evaluators for their cogent and useful evaluations. Used as convention, evaluation reports become bulk material for our files.

5. Evaluation as liturgy. When used in this way it doesn't matter what is said in the evaluation report, and the more arcane the language the better. It's enough to be able to say that an evaluation was done, that this or that eminent evaluation center was involved, and that the results are on file. We have made our bow to the alter of the evaluation gods. The liturgical use of evaluations is probably

related to the utility value of past evaluations experienced by the audience, or to the degree to which they believe they had better revere "evaluation."

There is a heirarchy suggested by the above uses to which clients put evaluations. The lower on that heirarchy our evaluations fall, the less likely it is that the field of evaluation is serving a purpose which will ensure its survival.

The Nature of What Is Being Evaluated

A qualitative evaluation of academic departments is quite different from evaluating the effectiveness of supplementary material developed for a particular course. Searching for the strengths and weaknesses in a curriculum is of a different order than helping an instructor evaluate his or her teaching with an eye toward improvement. The commnn element in each of the above roblems is simply that each can be evaluated. It is the nature of what is being evaluated which must determine methodologies, processes, and reporting formats of evaluation. To too many people evaluation still equals measurement. In spite of Scriven's (1967) clear statement that evaluation involves the estimation of merit, of worth, or of value, and in spite of Stake's (1969) encompassing statement that, "All evaluation deals explicitly with the worth of something," we have not been able to shake the belief that evaluation and measurement are synonomous terms.

Wardrop (1972) has delineated the concept of measurement from that of assessment on one hand and from evaluation on the other. By his reasoning, measurement basically involves the use of numerical values to represent attributes of objects. An attribute, in order to be measurable, must fit the specifications of a quantitative variable, and some unit of measurement must be established. Assessment includes measurement, but additionally involves those qualitative and judgemental activities which go into determining what and how to measure. His reasoning concludes that there are some aspects of evaluation which are outside the realm of both measurement and assessment. The

nature of what is being examined, and the purpose of the examination, are the determiners of whether one is involved in a measurement, assessment or evaluation activity. Wardrop's distinctions are worthy ones. An acceptance of his distinctions could further our efforts to develop a language for evaluation related more directly to its purposes.

There does seem to be a growing recognition that consumers are more and more demanding results from evaluation that have a real utility value rather than results that merely reflect the best efforts of a measurement technology. Flaugher (1974) is one of the more recent converts to a less rigid view of evaluation:

> " . . . educational evaluation properly done is a conglomeration of approximations to the ideal, obtrusive attempts to be unobtrusive, Kentucky windage, guesswork and wishful thinking. . . .
>
> "The most effective, useful evaluation is scrappy, rough, patchwork, frequently incomplete and crippled by unrealistic time deadlines which require short term guesses about reaching long term goals. It is different for every new media introduced: curriculum, text, film, computer assisted instruction."

Strange words indeed from a senior research psychologist with Educational Testing Service. Strange--but encouraging. House (1973) has gone so far as to suggest there is a "counterculture" in evaluation which has evolved in direct response to specific excesses and deficiencies of modern evaluation technology. And Provus (1973) has gone one step further in suggesting that, "A new professional is needed. A new system is needed which combines evaluation, community development, evaluation methodology, and public information into a unified whole."

The dissatisfaction with defining every evaluation problem as a measurement problem is plain. The recognition that the nature of the problem has an important bearing on the choice of method is an encouraging one.

LIST OF REFERENCES

Flaugher, Ronald, "Trends and Troubles in the Evaluation of Innovative Education." Paper delivered at the National Conference on Open Learning in Higher Education, Lincoln, Nebraska (January, 1974).

House, Ernest R., "Technology and Evaluation," Educational Technology, Vol. 13, No. 12 (1973).

Pell, Clairborne, "Building Partnerships for Educational Research and Development," Educational Researcher, Vol. 4, No. 1 (1975), pp. 11-12.

Provus, Malcolm, "In Search of Community," Phi Delta Kappan, Vol. 54, No. 10 (1973).

Scriven, Michael, "The Methodology of Evaluation." In Tyler, Gagne, and Scriven, Perspectives of Curriculum Evaluation. Chicago: Rand McNally and Co. (1967).

Stake, Robert, "Generalizability of Program Evaluation: The Need for Limits," EPIE Forum, Vol. 2, No. 5 (1969), pp. 39-41.

Stake, Robert, "Comments and Conjectures," Educational Forum, Vol. 36, No. 2 (1972[a]), pp. 271-272.

Stake, Robert, "An Approach to the Evaluation of Instructional Programs: Program Portrayal Versus Analysis." Paper delivered at the AERA annual meeting, Chicago (April 4, 1972[b]).

Stufflebeam, Daniel, Walter Foley, William Gephart, Egon Guba, Robert Hammond, Howard Merriman, and Malcolm Provus. Educational Evaluation and Decision Making, Itasca, Illinois: F. E. Peacock Publishers, Inc. (1971).

Wardrop, James L., "Toward a Broader Concept of Educational Assessment." Paper delivered at Second Annual Mardi Gras Symposium on Education: Instruction, Assessment, and Accountability, New Orleans (February ,1972).